

Northumbria Research Link

Citation: Teli, Shivani Rajendra, Zvanovec, Stanislav and Ghassemlooy, Zabih (2019) Performance evaluation of neural network assisted motion detection schemes implemented within indoor optical camera based communications. Optics Express, 27 (17). pp. 24082-24092. ISSN 1094-4087

Published by: Optical Society of America

URL: <https://doi.org/10.1364/OE.27.024082> <<https://doi.org/10.1364/OE.27.024082>>

This version was downloaded from Northumbria Research Link:
<http://nrl.northumbria.ac.uk/id/eprint/40275/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)



**Northumbria
University**
NEWCASTLE



UniversityLibrary

Performance evaluation of neural network assisted motion detection schemes implemented within indoor optical camera based communications

SHIVANI RAJENDRA TELI,^{1,*} STANISLAV ZVANOVEC,¹ AND ZABIH GHASSEMLOOY²

¹*Dept. of Electromagnetic Field, Faculty of Electrical Engineering, Czech Technical University in Prague, Prague, 16627, Czech Republic*

²*Optical Communications Research Group, Faculty of Engineering and Environment, Northumbria University, Newcastle-upon-Tyne, NE1 8ST, UK*

*telishiv@fel.cvut.cz

Abstract: This paper investigates the performance of the neural network (NN) assisted motion detection (MD) over an indoor optical camera communication (OCC) link. The proposed study is based on the performance evaluation of various NN training algorithms, which provide efficient and reliable MD functionality along with vision, illumination, data communications and sensing in indoor OCC. To evaluate the proposed scheme, we have carried out an experimental investigation of a static indoor downlink OCC link employing a mobile phone front camera as the receiver and an 8×8 red, green and blue light-emitting diodes array as the transmitter. In addition to data transmission, MD is achieved using a camera to observe user's finger movement in the form of centroids via the OCC link. The captured motion is applied to the NN and is evaluated for a number of MD schemes. The results show that, resilient backpropagation based NN offers the fastest convergence with a minimum error of 10^{-5} within the processing time window of 0.67 s and a success probability of 100 % for MD compared to other algorithms. We demonstrate that, the proposed system with motion offers a bit error rate which is below the forward error correction limit of 3.8×10^{-3} , over a transmission distance of 1.17 m.

© 2019 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

The optical wireless communications (OWC) technology covering ultraviolet, infrared and visible bands is a complementary technology to the dominated radio frequency (RF) based wireless systems that could be used for addressing the bandwidth bottleneck is a possible option for Internet of things (IoT) [1,2]. The visible spectrum band (i.e., 370-780 nm) known as visible light communications (VLC) is being considered as a possible option in 5th generation (5G) wireless networks for indoor environments. VLC utilizing the light-emitting diodes (LEDs) based lighting fixture offers four independent functionalities of data communications, illumination, localization and sensing in indoor environments [3,4]. In addition, VLC can offer massive MIMO (multiple in multiple output) capabilities using LED and photodetector (PD) arrays for IoT applications in both indoor and outdoor environments [5]. This feature of VLC is unique compared to massive MIMO in RF-based systems, which is too complex to implement.

The wide spread use of smartphones (six billion of them) with high-spec cameras are opening up new possibilities for VLC in applications where the need for high-data rate [6, 7]. Such applications include indoor localization, sensing, intelligent transportation systems, shopping areas, etc. The camera-based VLC, also termed as optical camera communications (OCC), has been studied within the framework of OWC and considered as part of the IEEE 802.15.7 standard [8,9]. OCC utilizes the built-in complementary metal-oxide-semiconductor camera

in smart devices as the receiver (Rx) for capturing two-dimensional data in the form of image sequences, thus enabling multidimensional data transmission. OCC with multiple functionalities of vision, data communications, localization and motion detection (MD) [8–10] can be used in all-optical IoT (OIoT) [5] based network application including device-to-device communications, mobile atto-cells, vehicle-to-everything (V2X), smart environments (home, office, surveillance), etc. [11]. In smart environments (i.e., home and offices) OCC-based MD can be utilized to effectively control smart devices [10, 12, 13]. This is very convenient and cost-effective as users carry smartphones with inbuilt cameras, which can be used as a Rx for both OCC and MD compared with other user interface methods such as gesture control (using a single webcam) and an infrared 3D camera for PC [14, 15]. MD based schemes such as (i) Li-Tech - shape detection and 3-D monitoring using visible light sensors [16]; (ii) Li-Sense - offering data communication and fine-grained, real-time human skeleton reconstruction using VL, which utilizes the shadowing effect and 324 PDs [17]; and (iii) a number of gesture recognition schemes [14, 15] have been proposed.

In OCC, image processing is critical for retrieving the transmitted data from the captured image frames. In recent years, intelligent machine-learning techniques (i.e., neural networks (NN)) have been adopted in image recognition for identifying objects' shape in the image, transcribing speech into a text, matching classified items and predicting the relevant results from network training [18]. In NN-based feature recognition schemes, multiple hidden layers with artificial neurons are used to train the network. These artificial neurons represent the main constituent, which receives multiple input samples in order to train the NN.

In [12] first reported by the authors, the initial results of MD performance based on images and centroid data samples (i.e., both considered as the input to NN representing the motion) using the variable learning rate backpropagation algorithm for training. The results in [12] demonstrate that, NN trained with centroid data samples performs only 5000 iterations in a time window of up to 4 s while the conventional NN trained using images can perform up to 8138 iterations in a time window of up to 9 s. Even though [12] provides a promising approach for MD, such long time windows could not be applied in real-time cases. Since the time windows have been obtained using a basic backpropagation algorithm, it is necessary to further perform more detailed analyzes of the proposed scheme based on the centroid data samples and using different transfer function-based algorithms for NN in order to reduce the time window and a number of iterations. In this paper, the focus is on the experimental investigation of NN-based MD for OCC performance using a number of transfer function-based training algorithms. In doing so, we include a wide scale of training parameters including the processing time (PT), iterations carried out by NN for MD, the percentage of success for MD and mean squared error (MSE). Unlike conventional NN schemes [18], the proposed NN-based MD is trained with centroid data samples and different transfer function algorithms, thus providing more accurate detection. In this work, experimental investigations are conducted for an indoor static downlink OCC with a smartphone front camera used as the Rx. The NN training is performed using eight different transfer function-based training algorithms for MD over the transmission distance L of up to 2 m. The OCC link quality in terms of the bit error rate (BER) and peak signal-to-noise ratio performance (PSNR) with respect to L is also analyzed simultaneously. The proposed NN-based MD can be used for control of data communications in OIoT networks.

The rest of the paper is structured as follows: Section 2 provides details of the proposed NN based MD in OCC. Experiment results are discussed in Section 3. Conclusions are drawn in Section 4.

2. Proposed NN based MD in OCC

2.1. System overview

Figure 1(a) illustrates the system overview of the proposed OCC-based NN assisted MD in an indoor environment. A data packet generation output, which is in a 12.8 kbits non-return-to-zero (NRZ) on-off keying (OOK) format, is first mapped according to the addresses of 8×8 red, green and blue (RGB) Neo pixel LED array using an Arduino Uno board (an open source microcontroller board based on the ATmega328 [19]). The intensity modulated (IM) light signal is transmitted over the free space channel. On the Rx side, an Android smartphone's front camera with the frame rate of 30 frames per second (fps) and a resolution of 1920×1080 pixels is used to capture the images (i.e., a video stream) of the IM LED array. In this work, the mobile phone is assumed to be located in a static position directly beneath the LED transmitter (Tx) at a height of 20 to 200 cm.

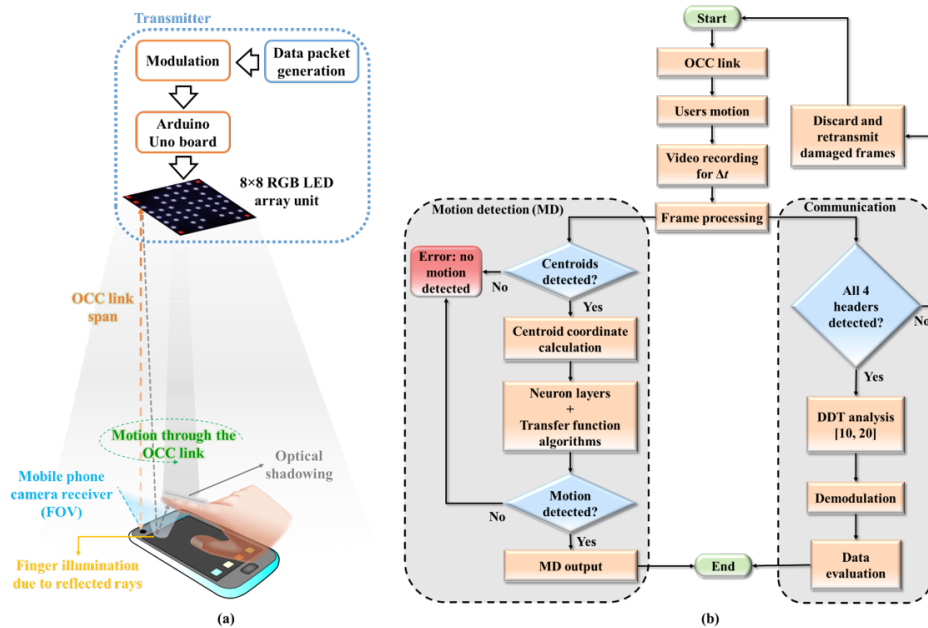


Fig. 1. System overview: (a) proposed NN assisted MD in OCC, and (b) a flowchart of MD and communication analysis.

Figure 1(b) shows the flowchart of MD and communication analysis on the Rx side. Note, motion is achieved by the user's finger moving over the camera. Both the RGB LEDs and finger movement are simultaneously captured by the camera in the form of a video stream, which is then divided into frames for post-image processing using MATLAB. Typically, the recorded video length depends on the motion duration Δt with the mean and maximum values of ~ 2.5 s and 4.5 s, respectively. For a camera, with a frame rate of 30 fps the captured frames of 75 and 135 are for Δt of 2.5 s and 4.5 s, respectively. As shown in Fig. 1(a), the user's finger hovering a few centimeters above the camera's screen will result in shadowing and reflected light rays. Note that, the illuminated finger is readily traceable by the camera using a tracking function, and its motion is expressed as centroids, which represent the center of a moving finger in the form of consecutive coordinate points [20], see Fig. 2. Note, in Fig. 2, each coordinate point represents the center of a moving finger tracked in a particular time frame. The key principle of MD is to compare the changes between the frames (a series of images) following video processing. The frame resolution is measured in terms of the pixels and inter-frame time, which is 33.3 ms (1 s/30 fps).

The motion between two consecutive frames can be simply determined as the difference between the centroid coordinates $(x_2 - x_1, y_2 - y_1)$ in $(N + 1)^{th}$ and N^{th} frame. The coordinate position of motion centroid (MC), which is obtained from the user's finger movement, is applied to a pre-trained NN system for detection and identification of user's motions. For the demonstration purpose, we consider five motion patterns, which are created from two simple natural motions in a straight, circular and curvature lines. These motions can be used to control smart devices, e.g., straight and circular motions can be used for turning ON and OFF of a device.

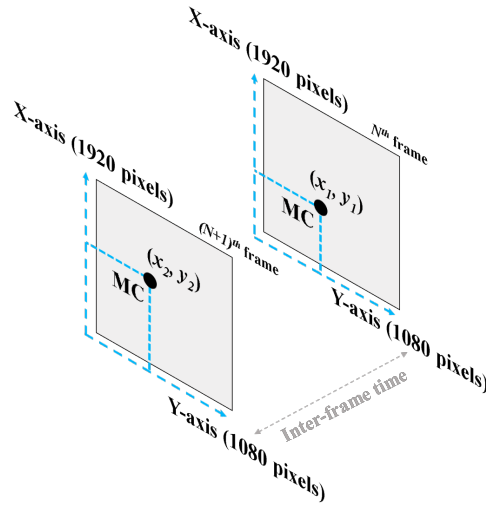


Fig. 2. Example of MC detection in two frames.

2.2. Data compensation scheme

We have adopted a transmit data compensation scheme based on the anchor LEDs (four per frame) and a synchronization LED for time synchronization, which is located in the first frame as in [12], in order to overcome blocking or shadowing due to mobility as depicted in Fig. 3(a). The data compensation scheme is based on discarding damaged frames due to the blocking of the anchor LEDs and requesting re-transmission. Note that, obstacles may fully/partially block one or more anchor LEDs, thus resulting in damaged frames, see Fig. 3(b), which will lead to increased BERs. The use of anchor LEDs (i.e., four-bit per frame in this case) results in reduced data throughput per frame, thus the trade-off between the BER and the data throughput.

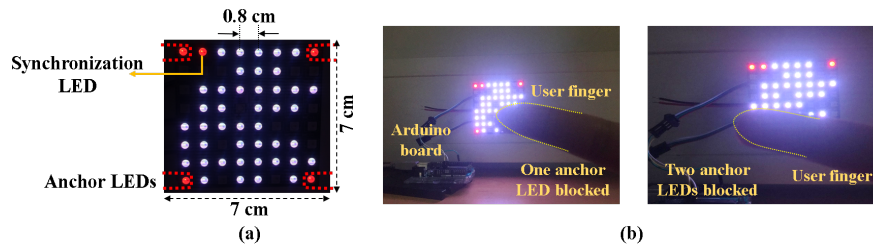


Fig. 3. (a) Transmitter configuration and (b) damaged frames due to blocking of one and two anchor LEDs.

For the proposed scheme with the transmit data compensation scheme, the data rate can be

given as:

$$R_d = N_L \times L_{FR} - N_A, \quad (1)$$

where N_L and N_A denote the number of data transmission, and anchors plus synchronization LEDs, respectively and L_{FR} is the flickering rate of LEDs (20 pulses per second in this work). For the proposed system, the maximum achievable R_d is 1.199 kbps (i.e., $64 \times 20 - 81$). Note, if the number of anchor LEDs is reduced the data throughput will slightly increase. E.g., for the anchor LEDs of 3, 2 and 1, the data rates are 1.219 kbps, 1.239 kbps and 1.259 kbps, respectively.

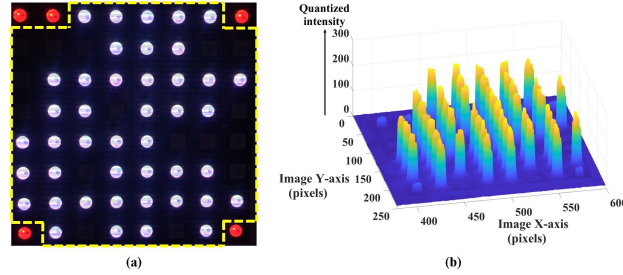


Fig. 4. (a) Identified data area excluding anchor and synchronization bits and (b) quantized intensity of the detected data.

For the proposed OCC-based scheme, we have adopted an efficient detection scheme of differential detection threshold (DDT) [10,21]. In the DDT scheme, the threshold level is defined in terms of the quantized intensity level within the range of [0-255]. Figure 4(a) represents the identified data area within the frame, while Fig. 4(b) provides the quantized intensity of the detected data. Based on DDT the initial value of threshold level was set to 181 level of quantized intensity as in [10,21]. Note, the threshold level can be adaptively set based on the intensity levels in the image frame.

2.3. NN-based MD for OCC

The proposed scheme can be trained using the transfer function algorithms in order to improve the MD performance by identifying only the predefined motions. Figure 5 illustrates the NN structure for MD performance evaluation within the context of OCC.

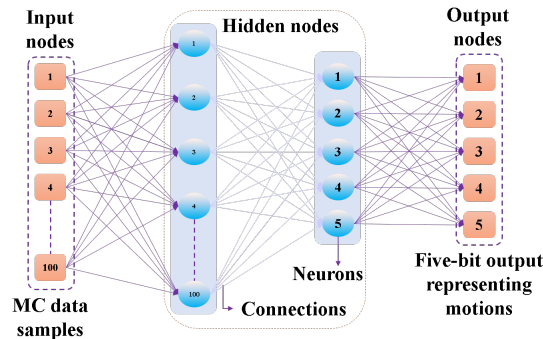


Fig. 5. NN structure for MD performance evaluation.

The input nodes are the coordinate positions of 100 centroid data samples, which represent 20 centroid data samples per predefined motions (i.e., variants of linear, circular and curvature

movements) for an OCC link span ranging from 20 cm to 200 cm. There are two hidden layers of 100 and 5 neurons, respectively. The hidden layers are used to detect and identify the user's motion, the output of which is expressed in the form of five-bit training labels representing the five predefined motions.

Note, for the training of NN, we have used eight possible transfer function-based algorithms as listed in Table 1. When selecting the most suitable training algorithm a number of factors needs considering including the number of neurons N_n in the hidden layers, PT, error measurement and the type of network used for pattern recognition, etc. [22]. In this work, we train the NN with MC and use pattern recognition to identify the classification of input signals or patterns in order to evaluate the link performance.

Table 1. **Training algorithms used for MD [21]**

Transfer function algorithm	Acronym
Levenberg-Marquardt	LM
Resilient Backpropagation	RP
Scaled Conjugate Gradient	SCG
Conjugate Gradient with Powell/Beale Restarts	CGB
Fletcher-Powell Conjugate Gradient	CGF
Polak-Ribiere Conjugate Gradient	CGP
One Step Secant	OSS
Gradient descent	GDX

The key system and NN training parameters are given in Table 2. The training parameters of training goal, iterations and time were set to the same values for all training algorithms in order to evaluate their performance under the same training environment.

Table 2. **System parameters.**

Description	Values
Capture device	Android mobile phone front camera with 1920×1080 pixels resolution and operated in the auto mode
Capture speed	30 fps
Tx	8 × 8 Neo pixel RGB LED array
Tx's flicker rate	20 Hz
Frame period at the Tx	50 ms
OCC link span	20 - 200 cm
NN hidden layers	two hidden layers with 100 and 5 neurons each
Epochs/training iterations	1000
Training time limit	infinite (allows NN to take sufficient time for training)

3. Results

Figures 6(a)-6(c) shows the experimental results of the detected MC representing variants of linear and circular motions as well as the curvatures. The solid grey line represents the actual considered motions while the dots represent detected MC tracked from user's finger movement over the smartphone's front camera when receiving data from the Tx. The coordinate points of these MC are used further to determine the probability of success for MD. Note, due to tracking some centroids are deviated from the actual motion path (highlighted in small blue circles) while some part of other light sources (highlighted in small red boxes) are captured within the surrounding. However, the NN training output is not affected due to these small number of deviated centroids and other light sources.

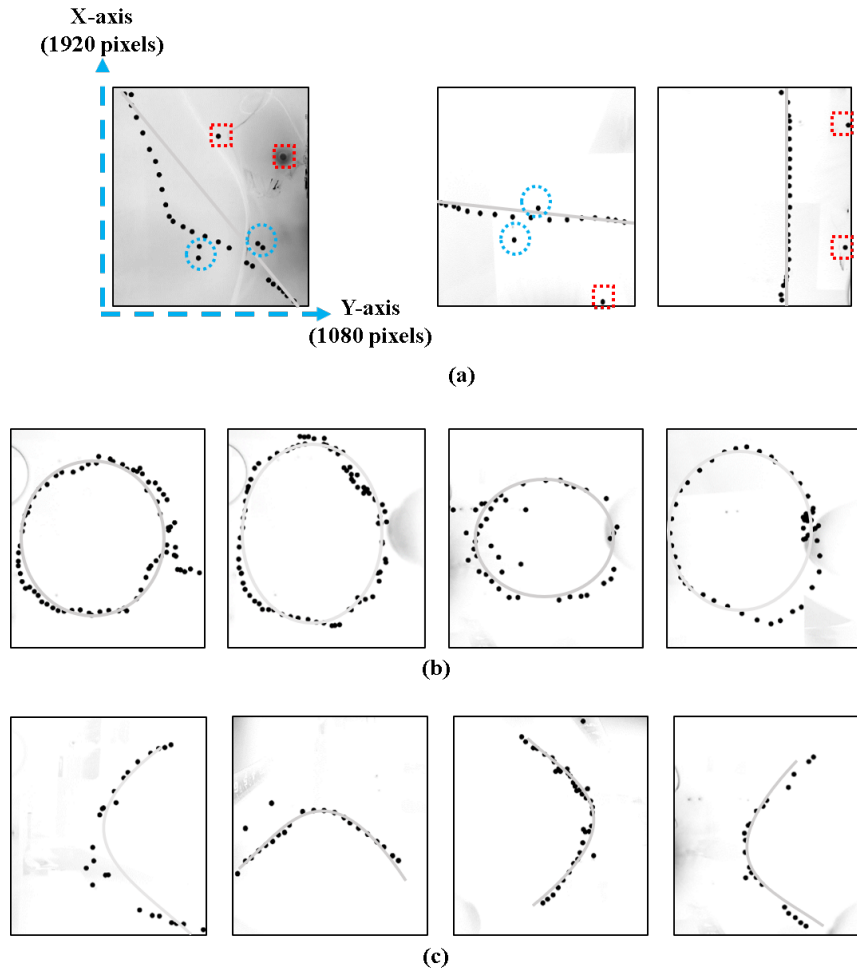


Fig. 6. Experiment results for MC representing variants of (a) linear motion, (b) circular motion, and (c) curvatures. Note, the solid grey line represents actual motion pattern while the dots represent MC.

To evaluate the system performance, we have used two criteria of MSE and the PT for all transfer function algorithms, which are obtained by averaging over 1000 training iterations for the OCC link span ranging from 20 - 200 cm, as depicted in Fig. 7(a). As mentioned in Table 2, the training time limit was set to infinite in order to examine properly all the training algorithms,

considering that some will take longer time to converge with the predicted accurate output. Note that, in a real-world application using NN with infinite networks a time complexity approach can be considered based on Markov Chain Monte Carlo method, which is compatible with large networks [23]. As shown in Fig. 7(a), the RP algorithm converges faster than others reaching the minimum MSE and PT of 5.1×10^{-5} and 0.67 s, respectively. The conjugate gradient algorithms (SCG, CGB, CGF and CGP) also perform well and can be used in networks with a large number of neuron weights due to the modest memory requirements [21]. Note, the LM algorithm offers the worst performance in terms of both the PT and MSE. This is because LM is designed for the least square problems, which are approximately linear in contrast to pattern recognition problems where the output neurons are generally saturated [21]. Both GDX and OSS algorithms converge rapidly provided the training is stopped too soon, but at the cost of inconsistent results [21]. Figure 7(b) illustrates the percentage of success for MD performed over a total of 100 experiments with respect to L for all algorithms listed in Table 1. The percentage of success for MD was determined based on the comparison of the exact input with five-bit output of NN, which represents the five predefined motions. It can be seen that RP display the best performance with the MD accuracies of 100 and 96.5 % over a link spans of 1.6 and 2 m (i.e., the maximum range in this work). The reduction in accuracy for increasing L is due to the fact that the illumination level of finger becomes lower as it moves away from the Tx. However, this does not have a significant impact on NN training and therefore, these reduced accuracy levels are still acceptable.

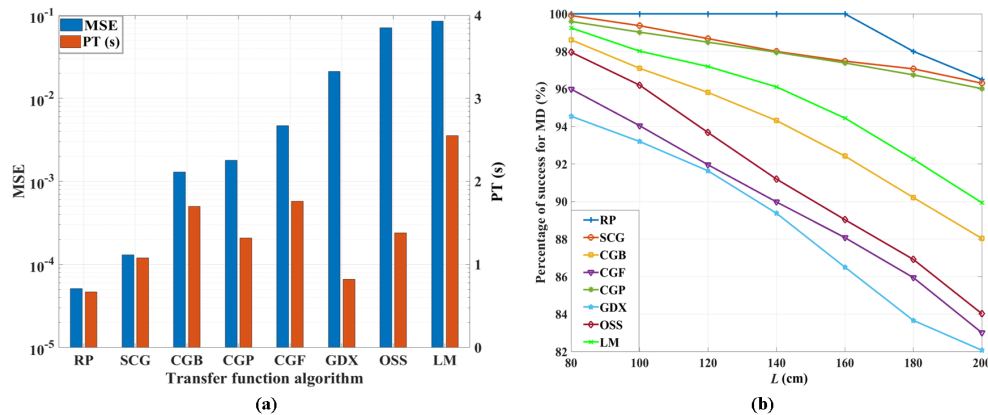


Fig. 7. MD performance: (a) the MSE and the PT, and (b) the percentage of success vs. the transmission distance L .

With RP displaying the best performance, we have further investigated its complexity of NN in terms of MSE and PT. Note, in general, N_n in the hidden layers can be larger or smaller than the number of input nodes (i.e., data samples). Large or small N_n will result in a complex NN and a higher number of training iterations and PT, respectively. Thus, the trade-off between N_n and NN training complexity is illustrated in Table 3.

Table 3. NN analysis.

NN layers	Epochs/training iterations	MSE	PT (s)
100-50-5-5	4000	2.1×10^{-3}	2 s
100-25-5-5	7000	4.3×10^{-2}	4.3 s
100-10-5-5	10000	5.3×10^{-2}	5.2 s

Since the proposed scheme offers simultaneous indoor data transmission via OCC and MD, next, we evaluated the link's BER and PSNR performances. Since in OCC the data is captured in the form of a two-dimensional image, a conventional SNR measurement cannot fully reflect the quality of the link. Therefore, we have adopted PSNR, which is widely used as a quality metric in image processing systems, as given by [24]:

$$\text{PSNR} = 10 \log_{10} \frac{I_{\text{peak}}^2}{\text{Avg}(I_{\text{Tx}} - I_{\text{Rx}})^2}, \quad (2)$$

where I_{peak}^2 denotes the squared peak intensity of the measured frame, I_{Tx} and I_{Rx} are the intensities of the transmitted and received frames.

Note that, user's motion will result in partial shadowing, which will ultimately affect the BER performance. Figure 8 shows the link's BER and PSNR performance against L for 12.8 kbits of data and four-bit header at a R_d of 1.199 kbps, where error-free data transmission is achieved at L up to 80 cm. Note, at the forward error correction (FEC) limit of 3.8×10^{-3} at L of 1.17 m, which is achieved because of the data compensation scheme. The transmission span of 1.17 m is a typical range in environments such as hospital wards, etc. Figure 8 depicts the BER performance as a function of PSNR for the proposed link. At a BER of 10^{-5} , well below the FEC limit of 3.8×10^{-3} , the PSNR is ~ 20 dB.

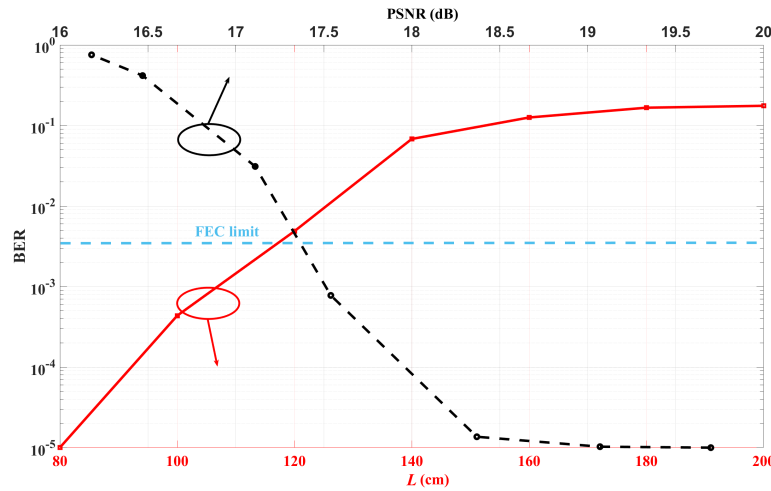


Fig. 8. OCC performance analysis: the BER performance as a function of the transmission span L , and the PSNR with respect to BER.

Finally, we compared the performance of the proposed NN assisted MD in OCC systems with MoC [10], TNMD [12], VLC-based MD [13] and LiSense [17] as shown in Table 4. In MoC, TNMD and the proposed NN assisted MD in OCC systems Android smartphone front camera has been used as Rx whereas, in VLC based MD and LiSense use PD-based Rx. Note, NN-based schemes offer improved performance compared to VLC-based systems. The highest percentage of success for MD of 96 % at L up to 200 cm is observed for the proposed scheme with the RP algorithm (with measured PSNR of 16.18 dB). The same percentage of success for MD is achieved for MoC a complex but a reliable Quadrant division based MD algorithm, but at L of 12 cm. Higher percentage of success for MD of 97 % is observed for TNMD with the basic NN algorithm at a maximum L of 125 cm. The improvement offered by the proposed NN assisted MD in OCC link, which uses a mobile phone camera as the Rx for MD and data transmission, is due to the use of RP algorithm within NN.

Table 4. Comparison of existing MD schemes.

	MoC [10]	TNMD [12]	VLC based MD [13]	LiSense [17]	NN assisted MD in OCC
Tx	8 × 8 red dot matrix LED array with 10 illumination LEDs	8 × 8 RGB LED array	20 RGB LEDs	5 off-the-shelf commercial LED lights	8 × 8 Neo pixel RGB LED array
Rx	Android smartphone front camera (1920 × 1080 pixels)	Android smartphone front camera (1920 × 1080 pixels)	9 PDs	324 PDs with 90° field of vision in a 3m × 3m area	Android smartphone front camera (1920 × 1080 pixels)
Detection algorithm	Quadrant division based MD algorithm	Trained neuron based MD	Pattern detection algorithm	LiSense: Sensing of light and analysing the shadow patterns	NN based detection using best training algorithm for MD
Motion distance (between detector and user)	7 - 8 cm	Up to 10 cm	Not specified	User moving in free-space between LEDs and PDs (2.65 m height)	Up to 10 cm
MD accuracy	96 % at L of 12 cm	97 % at L of 125 cm	Not specified but states high level of accuracy up to L of 60 cm	Blockage detection with respect to incident angle: 100 % at (0, 32) and 58-90 % at (43.2, 45)	96 % accuracy up to L of 200 cm
L	12 - 18 cm	25 - 200 cm	40 - 100 cm	Height of room: 265 cm	20 - 200 cm

4. Conclusion

The performance of NN assisted MD OCC link was experimentally evaluated for eight different transfer function based training algorithms with training parameters of PT, the number of iterations, the percentage of success for MD and MSE. We showed that, the best performance was achieved using the RP algorithm with the fastest convergence at a minimum error (MSE) and a PT of 10^{-5} and 0.67 s, respectively as well as the percentage of success for MD of 100 % up to a 1.6 m OCC link. For higher L , the OCC link will experience shadowing due to fingers' movement thus the need for diversity based Rx. We also demonstrated that, using the transmit data compensation scheme a high-quality data transmission with the FEC limit 3.8×10^{-3} , was achieved at 1.17 m OCC link. The reliability and efficiency of the proposed scheme were assessed

by comparing it with other existing techniques. The NN for MD analysis can be further extended to increase the link spans based on pattern recognition algorithms and using different transmitter configurations for mobility and multiuser indoor smart home environments. On the other hand, the data rate can be enhanced using a high capture speed camera with rolling shutter and a larger LED array as the Rx and the Tx, respectively in a MIMO OCC link.

Funding

H2020 Marie Skłodowska-Curie Innovative Training Network (VisIoN 764461).

References

1. P. H. Pathak, X. Feng, P. Hu, and P. Mohapatra, "Visible light communication, networking, and sensing: A survey, potential and challenges," *IEEE Commun. Surv. Tutorials* **17**, 2047–2077 (2015).
2. Z. Ghassemlooy, S. Zvanovec, M.-A. Khalighi, W. O. Popoola, and J. Perez, "Optical wireless communication systems," *Optik* **151**, 1–6 (2017).
3. S. Zvanovec, P. Chvojka, P. A. Haigh, and Z. Ghassemlooy, "Visible light communications towards 5G," *Radioengineering* **24**, 1–9 (2015).
4. Z. Ghassemlooy, W. Popoola, and S. Rajbhandari, *Optical wireless communications: system and channel modelling with Matlab®* (CRC press, 2019).
5. S. R. Teli, S. Zvanovec, and Z. Ghassemlooy, "Optical internet of things within 5G: Applications and challenges," in *2018 IEEE International Conference on Internet of Things and Intelligence System (IOTAIS)*, (IEEE, 2018), pp. 40–45.
6. R. Boubezari, H. Le Minh, Z. Ghassemlooy, and A. Bouridane, "Smartphone camera based visible light communication," *J. Light. Technol.* **34**, 4121–4127 (2016).
7. T. Nguyen, A. Islam, T. Hossan, and Y. M. Jang, "Current status and performance analysis of optical camera communication technologies for 5G networks," *IEEE Access* **5**, 4574–4594 (2017).
8. I. Takai, S. Ito, K. Yasutomi, K. Kagawa, M. Andoh, and S. Kawahito, "LED and CMOS image sensor based optical wireless communication system for automotive applications," *IEEE Photonics J.* **5**, 6801418 (2013).
9. M. J. Jang, "IEEE 802.15 WPAN 15.7 amendment-optical camera communications study group (SG 7a)," (2019 [Online accessed 6 March 2019]).
10. S. Teli, W. A. Cahyadi, and Y. H. Chung, "Optical camera communication: Motion over camera," *IEEE Commun. Mag.* **55**, 156–162 (2017).
11. Z. Ghassemlooy, L. N. Alves, S. Zvanovec, and M.-A. Khalighi, *Visible light communications: theory and applications* (CRC Press, 2017).
12. S. R. Teli, W. A. Cahyadi, and Y. H. Chung, "Trained neurons-based motion detection in optical camera communications," *Opt. Eng.* **57**, 1–4 (2018).
13. A. Sewaiwar, S. V. Tiwari, and Y.-H. Chung, "Visible light communication based motion detection," *Opt. Express* **23**, 18769–18776 (2015).
14. N. Lalithamani, "Gesture control using single camera for PC," *Procedia Comput. Sci.* **78**, 146–152 (2016).
15. D. Ionescu, V. Suse, C. Gadea, B. Solomon, B. Ionescu, and S. Islam, "A new infrared 3D camera for gesture control," in *2013 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, (IEEE, 2013), pp. 629–634.
16. E. Alizadeh Jarchlo, X. Tang, H. Doroud, V. P. G. Jimenez, B. Lin, P. Casari, and Z. Ghassemlooy, "Li-Tect: 3-D monitoring and shape detection using visible light sensors," *IEEE Sensors J.* **19**, 940–949 (2019).
17. T. Li, C. An, Z. Tian, A. T. Campbell, and X. Zhou, "Human sensing using visible light communication," in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, (ACM, 2015), MobiCom '15, pp. 331–344.
18. Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* **521**, 436–444 (2015).
19. Atmel datasheet, "8-Bit microcontroller with 4/8/16/32K bytes in-system programmable flash," (Atmel Corporation, 2009).
20. J. C. Nascimento, A. J. Abrantes, and J. S. Marques, "An algorithm for centroid-based tracking of moving objects," in *1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. ICASSP99 (Cat. No. 99CH36258)*, vol. 6 (IEEE, 1999), pp. 3305–3308.
21. Y. H. Kim and Y. H. Chung, "Experimental outdoor visible light data communication system using differential decision threshold with optical and color filters," *Opt. Eng.* **54**, 1–3 (2015).
22. Mathworks, "Choose a multilayer neural network training function," <https://in.mathworks.com/help/deeplearning/ug/choose-a-multilayer-neural-network-training-function.html;jsessionid=2e10b001e2c1c97fcac09f6004e6>.
23. C. K. I. Williams, "Computation with infinite neural networks," *Neural Comput.* **10**, 1203–1216 (1998).
24. Q. Huynh-Thu and M. Ghanbari, "Scope of validity of PSNR in image/video quality assessment," *Electron. Lett.* **44**, 800–801 (2008).